Australian Government
Department of Defence
Defence Science and
Technology Organisation

The Effect of Spectral Variation on Sound Localisation

Russell Martin, Ken McAnally, Tavis Watt and Patrick Flanagan

Air Operations Division
Defence Science and Technology Organisation

DSTO-RR-0308

ABSTRACT

Previous research has suggested that spatialised auditory displays will enhance operator performance in many military settings. It is well known that a sound's spectrum must be broad and relatively flat for the sound to be accurately localised. The study described here examined the effect of systematically varying the evenness of a sound's spectrum on the accuracy with which the sound can be localised. Six participants localised spectrally scrambled sounds produced by setting the sound levels in the 98-, 391- or 1562-Hz wide frequency bands comprising a broadband (0-25 kHz) sound to random values within a 0-, 20-, 40- or 60-dB range. Localisation errors were found to increase with increases in both bandwidth and band-level range. Scrambling the spectra of sounds over a 60 dB range led to as much as a doubling of mean elevation error and a trebling of front/back confusion rate. The accuracy with which these sounds could be localised was found to be highly correlated with a simple measure of spectral variation. The results of this study inform the development of guidelines for designing localisable sounds to be used in spatialised auditory displays.

RELEASE LIMITATION

Approved for public release

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAR 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	
The Effect of Spectral Variation on Sound Location	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **DSTO Defence Science and Technology Organisation 506 Lorimer Street Fishermans Bend, Victoria 3207 Australia**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**25**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Published by

Air Operations Division
DSTO Defence Science and Technology Organisation
506 Lorimer St
Fishermans Bend, Victoria 3207 Australia

Telephone: (03) 9626 7000
Fax: (03) 9626 7999

AR-013-607
March 2006

APPROVED FOR PUBLIC RELEASE

The Effect of Spectral Variation on Sound Localisation

Executive Summary

Auditory displays are used in a variety of military settings. In many of these settings it would be advantageous if information concerning the locations of particular entities and/or events could be conveyed. For example, spatialised auditory threat warnings could greatly assist operators of military aircraft, who often need to know a threat's location in order to respond to it appropriately.

Recently developed three-dimensional (3D) audio technology provides a means of conveying spatial information via auditory displays. For a 3D auditory display to impart useful spatial information the sounds presented through it must be accurately localised. The study described here is a step in the process of developing guidelines for the design of localisable sounds intended for use with 3D auditory displays.

Previous research has indicated that a sound's spectrum must be broad and relatively flat for the sound to be accurately localised. In this study, the effect of systematically varying the evenness of a sound's spectrum on the accuracy with which the sound can be localised was examined. Participants in the study localised spectrally scrambled sounds that were produced by setting the sound levels in the 98-, 391- or 1562-Hz wide frequency bands making up a broadband sound to random values within a range of 0-, 20-, 40- or 60-dB. It was found that the accuracy with which these sounds could be localised decreased with increases in both the width of the frequency bands and the range of the sound-level variation. Scrambling the spectra of sounds over a 60 dB range resulted in the mean elevation error increasing by a factor of up to two and the rate at which front and rear sound-source locations are confused increasing by a factor of up to three.

Of particular interest, it was found that the accuracy with which the sounds in this study could be localised was highly correlated with a simple measure of the variation in a sound's spectrum. Further research is required to ascertain whether this measure of spectral variation is predictive of the accuracy with which a wider variety of sounds can be localised.

Authors

Russell Martin
Air Operations Division

Russell Martin is a Senior Research Scientist in Human Factors in Air Operations Division. He received a Ph.D. in Psychology from Monash University in 1988 and worked at the University of Queensland, Oxford University, the University of Melbourne and Deakin University prior to joining DSTO in 1995.

Ken McAnally
Air Operations Division

Ken McAnally is a Senior Research Scientist in Human Factors in Air Operations Division. He received a Ph.D. in Physiology and Pharmacology from the University of Queensland in 1990 and worked at the University of Melbourne, the University of Bordeaux and Oxford University prior to joining DSTO in 1996.

Tavis Watt
Deakin University

At the time the research described here was conducted, Tavis Watt was an honours student in the School of Psychology at Deakin University. His current position is: Operations Manager, Property Services Division, Deakin University.

Patrick Flanagan
Deakin University

Patrick Flanagan is a Lecturer in the School of Psychology at Deakin University. He received a Ph.D. in Psychology from Queens University in 1987.

Contents

1. Introduction

Auditory displays are used in a variety of military settings to convey critical information. In many of these settings it would be advantageous if information concerning the spatial locations of particular entities and/or events could be conveyed. For example, several studies [McKinley, Erikson & D'Angelo 1994; Perrott, *et al.* 1996; Bronkhorst, Veltman & van Breda 1996; Nelson *et al.* 1998; Bolia, D'Angelo & McKinley 1999; Parker *et al.* 2004] have suggested that spatialised threat warnings would be of considerable benefit to operators of military aircraft, who often need to know a threat's location in order to respond to it appropriately.

With the development of three-dimensional (3D) audio technology it has become possible to generate headphone-presented sounds that listeners perceive to come from remote sources at distinct locations in space [e.g., Wightman & Kistler 1989; Bronkhorst 1995; Carlile *et al.* 1996; Martin, McAnally & Senova 2001]. The application of this technology in conjunction with warning systems has substantial potential. For a 3D audio display to impart useful spatial information the sounds presented through it must be accurately localised.

It is well known that the accuracy with which a sound can be localised is dependent on its spectral content. Many studies [e.g., Roffler & Butler 1968a; Hebrank & Wright 1974; Butler & Planert 1976; King & Oldfield 1997] have demonstrated that localisation accuracy diminishes as the bandwidth of a sound is systematically reduced. Localisation judgments for sounds having particularly narrow bandwidths, such as tones and narrow-band noises, can be influenced more by the sound's centre frequency than by the location of its source [Pratt 1930; Roffler & Butler 1968b; Blauert 1969/70; Musicant & Butler 1985; Middlebrooks 1992]. Vertical and front/back components of localisation judgments are particularly vulnerable in this respect [e.g., Middlebrooks 1992]. It appears that a sound's spectrum must extend from about 1 to 16 kHz for localisation to be optimal [Hebrank & Wright 1974; King & Oldfield 1997].

The requirement of broad sound bandwidth for accurate localisation stems in part from the role of spectral cues in the localisation process. Spectral cues, which result from the interaction of sound with the torso, head and pinnae, are believed to provide information that can resolve the ambiguity inherent in interaural time and level difference cues to a sound's location [see Middlebrooks & Green 1991, for a review]. Analyses of human free-field to eardrum transfer functions have revealed the presence of features, such as prominent peaks or notches in the transfer functions' magnitude spectra, that vary with source location [Shaw & Teranishi 1968; Blauert 1969/70; Hebrank & Wright 1974; Mehrgardt & Mellert 1977]. It is thought that listeners learn to associate particular patterns of spectral features in the signals at their ears with particular sound-source locations. For locations well-removed from the median vertical plane, the spectral cues provided by the near ear completely dominate those provided by the far ear [e.g., Humanski & Butler 1988; Morimoto 2001]. For locations in the vicinity of the median plane, the cues provided by both ears contribute to perceived source location [Morimoto 2001; Hofman & Van Opstal 2002].

For the pattern of spectral features in the signal at a listener's ear to provide a valid localisation cue, the listener must make a correct assumption about the spectrum of the sound at its source (i.e., the presence of a particular feature in a signal at an ear could reflect the imposition of that feature on an incoming sound by the torso, head and/or pinna, or the presence of that feature in the sound at its source). It has been found that listeners have considerable difficulty distinguishing at-ear spectral features that result from location-dependent filtering from those that reflect the spectrum of the sound at its source [see, for example, Rakerd, Hartmann & McCaskey 1999]. Studies in which localisation has been found to be disrupted when a narrow frequency band is removed from a broadband stimulus [Hebrank and Wright 1974; Butler & Musicant 1993; Burlinghame & Butler 1998] or when sound levels in the narrow frequency bands comprising broadband stimuli are varied randomly [Wightman & Kistler 1997] suggest that listeners usually assume that the spectrum of a sound at its source is relatively flat.

The fidelity of a 3D audio display can be expected to be optimal, therefore, when the sounds presented through it have relatively flat spectra. But exactly how flat is relatively flat? Wightman and Kistler's [1997] "scrambled-spectrum" stimuli were produced by setting the sound level in each critical band (a frequency band about 1/6 of an octave wide) of a broadband stimulus to a random value within a 20- or 40-dB range. The effect on localisation of scrambling the spectra of stimuli over a 40-dB range was reported by Wightman and Kistler to differ across listeners. For one "typical" listener the effect described was a reduction in the accuracy with which sound-source elevation could be discerned and an increase in the incidence of front/back confusions (i.e., occasions on which the sound source was judged to be in the incorrect front-versus-back hemifield). The effect on localisation of scrambling the spectra of stimuli over a 20-dB range is difficult to determine from the data presented by Wightman and Kistler but appears to be less pronounced than that of scrambling the spectra of stimuli over a 40-dB range.

As Wightman and Kistler's [1997] study was not primarily concerned with the effect of spectral variation on sound localisation, the width of the frequency bands in which sound levels were randomised was not varied. It is likely, however, that variation in a sound's spectrum will have a greater effect on the accuracy with which the sound can be localised when the coarseness of that variation more closely resembles that of the spectral cues used by listeners to localise sound. In the study described here, the effect on localisation of varying a sound's spectrum was examined using scrambled-spectrum stimuli produced by setting the sound level in each of the 98-, 391- or 1562-Hz wide frequency bands comprising a broadband stimulus to a random value within a 20-, 40- or 60-dB range. The results of this study inform the development of guidelines for designing localisable sounds to be used in spatialised auditory displays.

2. METHODS

2.1 Participants

Six volunteers, one female and five male, ranging in age from 22 to 39 years participated in this study. The hearing of each was assessed by measuring his or her absolute thresholds for 1, 2, 4, 8, 10, 12, 14 and 16 kHz pure tones using procedures described in detail by Watson et al. [2000]. For all participants all thresholds were lower than the relevant age-specific norm [Corso 1963; Stelmachowicz *et al.* 1989].

Each participant was allowed to practice localising broadband (0.05-20 kHz) noise during several training sessions prior to data collection. All participants demonstrated a high level of proficiency (a mean localisation error of less than 14°) at this task.

2.2 Stimulus generation and presentation

On each trial an independent sample of Gaussian noise (328 ms in duration and incorporating 20-ms cosine-shaped rise and fall times) was generated at a sampling rate of 50 kHz (Tucker-Davis Technologies AP2) and passed through a broadband (0-25 kHz) digital filter designed and implemented in the frequency domain. A new filter was constructed for each trial by dividing the frequency region extending from 0 to 25 kHz into 98-, 391- or 1562-kHz wide bands and setting the level in each band to a constant value (0-dB band-level range) or a random value within a range of 20, 40 or 60 dB. Where the band-level range was 20, 40 or 60 dB, the range of levels in any given filter tended to be somewhat less than the full available range. This was particularly the case for the broadest bandwidth. Mean level-ranges for filters generated by randomising sound levels in 98-, 391- or 1562-Hz wide bands within 20-, 40- or 60-dB ranges are shown in Table 1. The resulting filtered noise was passed through a second digital filter to compensate for the transfer characteristics of the stimulus presentation system.

Table 1: Mean level-ranges for filters generated by randomising sound levels in 98-, 391- or 1562-Hz wide bands within 20-, 40- or 60-dB ranges. Each value is the mean of the ranges of 1000 filters.

Band-level range (dB)	**Bandwidth (Hz)**		
	98	***391***	***1562***
20	19.8	19.4	17.6
40	39.7	38.8	35.3
60	59.5	58.2	52.9

Stimuli were converted to analogue signals (Tucker-Davis Technologies PD1), passed through an anti-aliasing filter with a low-pass cut-off frequency of 20 kHz (Tucker-Davies Technologies FT5), amplified (Hafler Pro 1200) and presented at 65 dB SPL (A-weighted)

through a loudspeaker (Bose Free Space) mounted on a 1-m radius hoop. Loudspeaker movement was driven by programmable stepping motors that could position the loudspeaker at any location from 0 to 359.9° azimuth and –50 to +80° elevation with a resolution of 0.1°. A convention of measuring azimuth in the clockwise direction and describing elevation below the interaural horizontal plane as negative was followed.

2.3 Procedure

The participant was seated in a sound-attenuated anechoic chamber such that his or her head was positioned at the centre of the loudspeaker hoop. Background noise levels within the chamber were less than 10 dB SPL in all 1/3-octave bands with centre frequencies ranging from 0.5 to 16.0 kHz. The participant's view of the hoop and loudspeaker was obscured by a cloth sphere supported by thin fibreglass rods. The cloth from which this sphere was constructed was acoustically transparent. A dim light inside the sphere allowed visual orientation. Participants wore a headband on which a laser pointer and a magnetic tracker receiver (3 Space Fastrak, Polhemus) were mounted.

At the beginning of each trial the participant fixated on a light emitting diode (LED) positioned at 0° azimuth and elevation. When ready, he or she pressed a hand-held button. The loudspeaker was then moved to the target location. Loudspeaker movement occurred in two steps to reduce the likelihood of participants discerning the target location from the duration of movement. During the first step the loudspeaker was moved to a randomly chosen location at least 30° distant in both azimuth and elevation from the previous and subsequent target locations. During the second it was moved to the target location. (In tests conducted previously to the experiment described here it was found that the accuracy with which participants could discern the target location in the absence of an acoustic stimulus was no greater than that expected on the basis of chance.) The target location for each trial was chosen randomly from the set ranging from 0 to 359.9° azimuth and –45 to +75° elevation in 0.1° steps. The location selection algorithm ensured that locations were distributed more-or-less evenly across this part-sphere by ensuring that extreme elevations were not overrepresented. (The probability of any given elevation being selected was proportional to the circumference of the sphere at that elevation.) As soon as the loudspeaker was in position, the LED was turned off. The LED then flashed three times to alert the participant. The acoustic stimulus was presented immediately thereafter. The participant was requested to keep his or her head stationary during presentation of the stimulus.

Following stimulus presentation, the head-mounted laser pointer was turned on and the participant directed the laser toward the precise point on the surface of the cloth sphere from which he or she perceived the stimulus to come. The location and orientation of the laser pointer were measured using the magnetic tracker, and the point where the beam intersected the sphere was calculated geometrically. An LED attached to the centre of the loudspeaker was then turned on and the participant directed the laser toward the LED. The location and orientation of the laser pointer were measured again and the point where the beam intersected the sphere was calculated geometrically.

Localisation accuracy was described in terms of two errors: lateral error and elevation error. Lateral error was defined as the unsigned difference between the true and perceived sound-source lateral angles, where lateral angle is the angle subtended at the hoop centre between the sound source and the vertical plane separating the left and right hemispheres of the hoop. Elevation error was defined as the unsigned difference between the true and perceived sound-source elevations, where elevation is the angle subtended at the centre of the hoop between the sound source and the horizontal plane separating the upper and lower hemispheres of the hoop.

Each experimental session contained 40 trials and involved a single combination of bandwidth (98, 391 or 1562 Hz) and band-level range (0, 20, 40 or 60 dB). Each participant took part in 36 sessions, 3 at each of the 12 possible bandwidth and band-level range combinations. The order in which these 12 conditions were presented was determined following a randomised-blocks procedure. Participants completed a maximum of two sessions per day.

For each participant a mean lateral and elevation error was calculated for each condition after data for those trials on which a front/back confusion was made had been removed. A front/back confusion was deemed to have been made if two conditions were met. The first was that neither the true nor the perceived sound-source location fall within a narrow exclusion zone symmetrical about the vertical plane dividing the front and back hemispheres of the hoop. The width of this exclusion zone, in degrees of azimuth, was 15 divided by the cosine of the elevation. (Note that the arc length associated with 1° of azimuth is greatest at 0° of elevation and becomes progressively smaller as either vertical pole is approached.) The second condition was that the true and perceived sound-source locations be in different front-versus-back hemispheres. The proportion of front/back confusions was calculated for each participant and condition by dividing the number of trials on which a front/back confusion was made by the number of trials on which neither the true nor the perceived sound-source location fell within the exclusion zone.

Data were analysed using two-way repeated-measures analyses of variance incorporating Greenhouse-Geisser corrections for violations of the assumption of sphericity where appropriate [Keppel 1991]. The *a priori* alpha level was set at 0.05.

3. RESULTS

The mean lateral error averaged across participants is shown in Figure 1 for each of the 12 bandwidth and band-level range combinations. The mean lateral error varied little across these conditions and ranged from 5.3 to 6.5°. Statistical analysis indicated that neither the main effect of bandwidth nor that of band-level range was significant (bandwidth: $F(1.6,8.2)=3.46$, $p=0.087$; band-level range: $F(1.4,6.9)=4.48$, $p=0.066$). The interaction between bandwidth and band-level range was also found not to be significant ($F(2.6,12.9)=2.23$, $p=0.140$).

The mean elevation error averaged across participants is shown in Figure 2 for each of the bandwidth and band-level range combinations. The mean elevation error increased notably with both increasing bandwidth and increasing band-level range. The extent to which it increased with increasing band-level range was greater for broader bandwidths. For bandwidths of 98, 391 and 1562 Hz it increased from 6.2 to 9.4°, 6.7 to 12.7° and 6.3 to 14.3°, respectively. Statistical analysis indicated that the main effects of bandwidth and band-level range and the interaction between these variables were significant (bandwidth: $F(1.3,6.3)=33.27$, $p<0.001$; band-level range: $F(2.5,12.3)=76.20$, $p<0.001$; interaction: $F(2.5,12.6)=5.38$, $p=0.016$). Planned comparisons revealed that the mean elevation error was significantly greater for the 40 and 60 dB band-level ranges compared with the 0 dB band-level range for the 98-Hz bandwidth (0 vs. 40 dB: $F(1,5)=24.04$, $p=0.004$; 0 vs. 60 dB: $F(1,5)=34.17$, $p=0.002$), the 20, 40 and 60 dB band-level ranges compared with the 0 dB band-level range for the 391-Hz bandwidth (0 vs. 20 dB: $F(1,5)=13.41$, $p=0.015$; 0 vs. 40 dB: $F(1,5)=21.2$, $p=0.006$; 0 vs. 60 dB: $F(1,5)=161.32$, $p<0.001$) and the 20, 40 and 60 dB band-level ranges compared with the 0 dB band-level range for the 1562-Hz bandwidth (0 vs. 20 dB: $F(1,5)=16.62$, $p=0.010$; 0 vs. 40 dB: $F(1,5)=45.36$, $p=0.001$; 0 vs. 60 dB: $F(1,5)=229.58$, $p<0.001$).

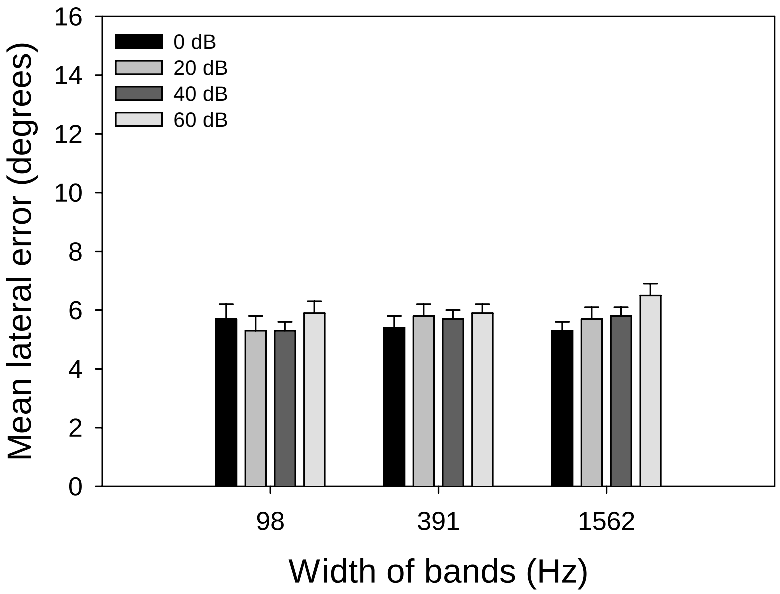

Figure 1: Mean lateral error averaged across participants for each of the three bandwidths and four band-level ranges. Each error bar shows one standard error of the average.

The proportion of front/back confusions averaged across participants is shown in Figure 3 for each of the bandwidth and band-level range combinations. As was the case for the mean elevation error, the proportion of front/back confusions tended to increase with increasing bandwidth and band-level range. The extent to which the proportion of front/back confusions increased across band-level ranges was greater for the 391- and 1562-Hz bandwidths than for the 98-Hz bandwidth. Statistical analysis indicated that the main effects of bandwidth and band-level range were significant (bandwidth: $F(1.7,8.4)=18.76$, $p=0.001$; band-level range: $F(1.8,9.1)=16.51$, $p=0.001$) but the interaction between these variables was not significant ($F(3.2,15.6)=1.62$, $p=0.223$). Planned comparisons revealed that the proportion of front/back confusions was significantly

greater for the 40 and 60 dB band-level ranges compared with the 0 dB band-level range for the 98-Hz bandwidth (0 vs. 40 dB: $F(1,5)=6.86$, $p=0.047$; 0 vs. 60 dB: $F(1,5)=26.50$, $p=0.004$), the 20, 40 and 60 dB band-level ranges compared with the 0 dB band-level range for the 391-Hz bandwidth (0 vs. 20 dB: $F(1,5)=11.51$, $p=0.019$; 0 vs. 40 dB: $F(1,5)=10.16$, $p=0.024$; 0 vs. 60 dB: $F(1,5)=36.26$, $p=0.002$) and the 20, 40 and 60 dB band-level ranges compared with the 0 dB band-level range for the 1562-Hz bandwidth (0 vs. 20 dB: $F(1,5)=9.93$, $p=0.025$; 0 vs. 40 dB: $F(1,5)=36.82$, $p=0.002$; 0 vs. 60 dB: $F(1,5)=11.62$, $p=0.019$).

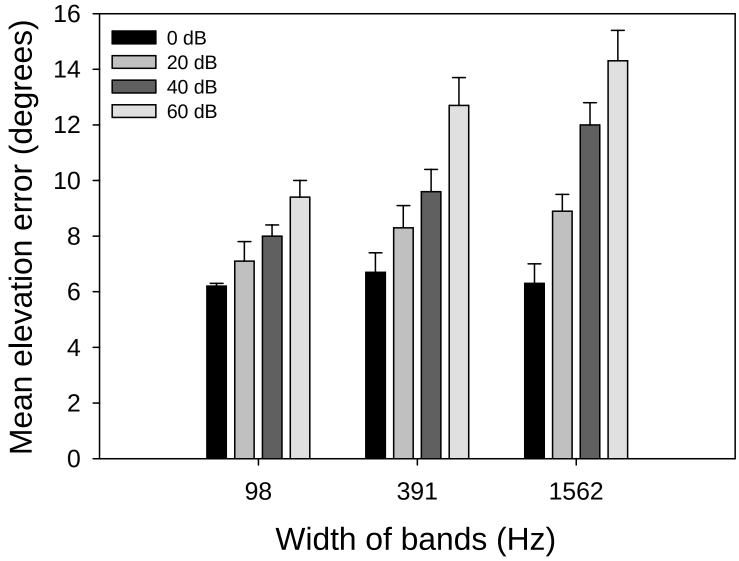

Figure 2: Mean elevation error averaged across participants for each of the three bandwidths and four band-level ranges. Each error bar shows one standard error of the average.

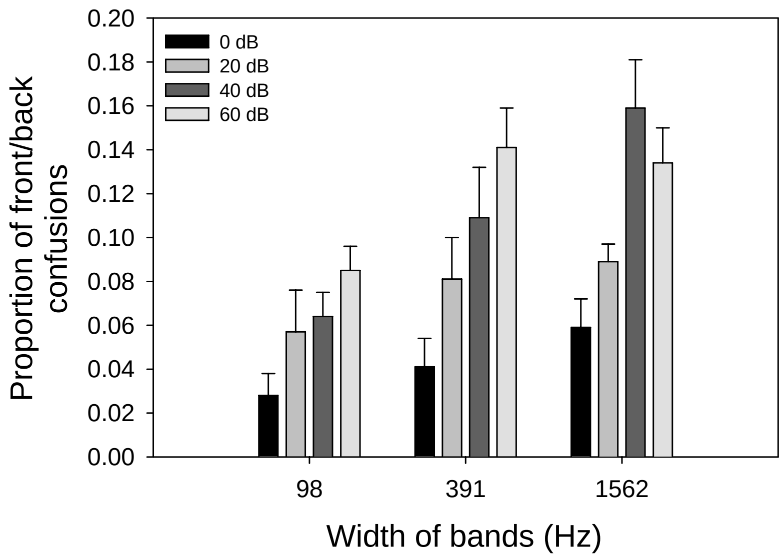

Figure 3: Mean proportion of front/back confusions averaged across participants for each of the three bandwidths and four band-level ranges. Each error bar shows one standard error of the average.

That the effects of spectral scrambling on the accuracy of judgments of sound-source elevation and front/back hemifield were greater for stimuli generated by randomising sound levels in broader spectral bands is likely to have resulted from the spectral smoothing imposed on stimuli by the cochlea (see Discussion section for a more detailed argument). This smoothing can be approximated by passing stimuli through a set of 1/3-octave bandpass filters. Mean level-ranges across the eight 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz (and therefore encompassing the portion of the spectrum believed to contain the most important spectral cues) for the spectrally scrambled stimuli in this study are shown in Table 2. It can be seen that mean level-range following 1/3-octave band filtering increases with increasing bandwidth for each band-level range.

Table 2: Mean level-ranges across the eight 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz for stimuli generated by randomising sound levels in 98-, 391- or 1562-Hz wide bands within 20-, 40- or 60-dB ranges. Each value is the mean of the ranges of 1000 stimuli.

Band-level range (dB)	**Bandwidth (Hz)**		
	98	***391***	***1562***
20	8.6	11.3	14.2
40	10.8	17.3	24.5
60	13.2	22.0	31.6

Level range across 1/3-octave bands is one simple measure of the spectral variation in an acoustic signal. Another measure, which we have found to be more predictive of localisation accuracy, is the sum of the absolute differences between levels in adjacent 1/3-octave bands. In Figures 4 and 5, the mean sum of differences between levels in adjacent 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz is plotted against the mean elevation error and proportion of front/back confusions, respectively, for each of the 12 combinations of bandwidth and band-level range. Dashed lines show the linear regressions of mean elevation error and proportion of front/back confusions on the mean sum of differences. It can be seen that the sum of differences between levels in adjacent 1/3-octave bands is highly predictive of both the mean elevation error (in which case r^2=0.97) and the proportion of front/back confusions (in which case r^2=0.89).

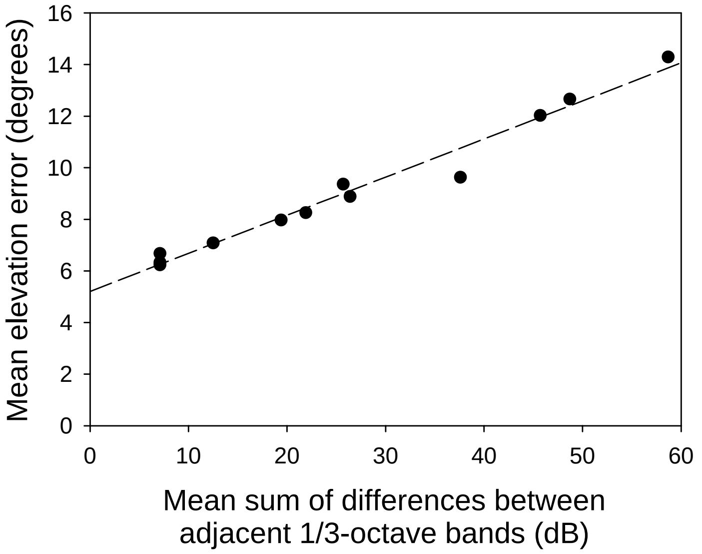

Figure 4: Mean sum of differences between levels in adjacent 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz plotted against the mean elevation error averaged across participants for each of the 12 combinations of bandwidth and band-level range. Each mean sum of differences is the mean of the sums of differences of 1000 stimuli. The dashed line shows the linear regression of mean elevation error on mean sum of differences between adjacent 1/3-octave bands. r^2=0.97.

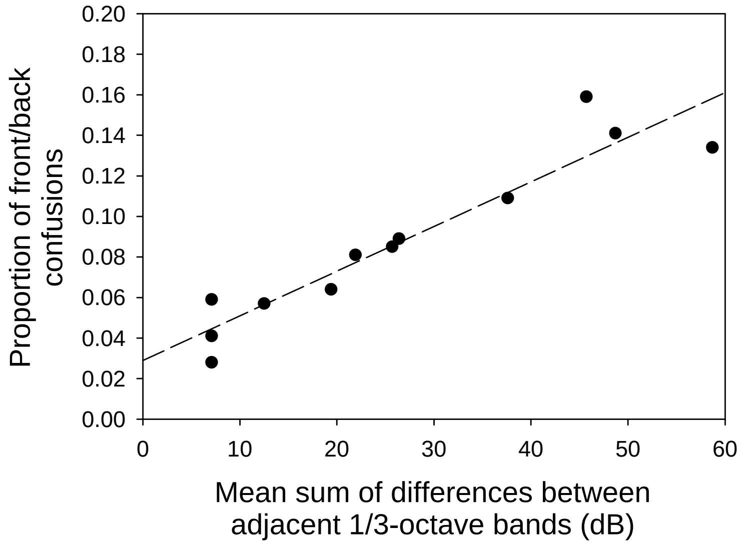

Figure 5: Mean sum of differences between levels in adjacent 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz plotted against the proportion of front/back confusions averaged across participants for each of the 12 combinations of bandwidth and band-level range. Each mean sum of differences is the mean of the sums of differences of 1000 stimuli. The dashed line shows the linear regressions of proportion of front/back confusions on mean sum of differences between adjacent 1/3-octave bands. r^2=0.89.

4. DISCUSSION

The results of this study indicate that scrambling the spectra of otherwise flat broadband sounds over a range as small as 20 dB can significantly reduce the accuracy with which the stimuli are localised. In proportional terms, this manipulation seems to have its greatest effect on front/back confusion rates, which increased by a factor of about 2 when band-level range was increased from 0 to 20 dB. Scrambling the spectra of stimuli over a 40 or 60 dB range was found to have a greater effect, and led to as much as a doubling of mean elevation error (e.g., 0 vs. 60 dB band-level ranges, 1562 Hz bandwidth) and a trebling of front/back confusion rate (e.g., 0 vs. 60 dB band-level ranges, 391 Hz bandwidth). It is worth noting, however, that the spectral scrambling employed in this study, which arguably was quite severe, was not completely disruptive of localisation performance. For example, the largest mean elevation error of 14.3° for the 1562-Hz bandwidth/60-dB band-level range condition is considerably smaller than the mean elevation error of 33.3° that would be expected (on the basis of a simulation involving 10^6 trials) if participants perceived the lateral angle of the sound source accurately but responded randomly with respect to its elevation. Likewise, the largest front/back confusion proportion of 0.159 for the 1562-Hz bandwidth/40-dB band-level range condition is considerably smaller than the front/back confusion proportion of 0.5 that would be expected if participants responded randomly. Similarly moderate effects of extreme spectral scrambling on localisation accuracy are evident in the data presented by Wightman and Kistler [1997].

The results of this study are also consistent with those presented by Macpherson and Middlebrooks [2003], who examined the accuracy with which stimuli with rippled (sinusoidally on a logarithmic frequency scale) amplitude spectra of differing ripple densities and depths can be localised. Macpherson and Middlebrooks found that localisation was most disrupted, relative to that for flat-spectrum stimuli, for ripple densities ranging from 0.5 to 2 ripples/octave (when ripple depth was held constant at 40 dB) and ripple depths of at least 20 dB (when ripple density was held constant at 1 ripple/octave). Even under these conditions, however, the observed disruption was only moderate. Spectral ripples having depths of less than 20 dB were found to have little effect on localisation.

It is clear that the effect of spectral scrambling is to increase elevation errors and front/back confusion rates while leaving lateral errors generally unaltered. This is consistent with the expectation that spectral-scrambling would disrupt spectral cues to sound-source location but not affect interaural time or level difference cues. These latter cues are thought to indicate the 'cone-of-confusion' on which a sound source lies [Mills 1972], which defines the source's lateral angle. Spectral cues, on the other hand, are thought to resolve the ambiguity inherent in interaural time and level difference cues by indicating the source's position around the cone of confusion (i.e., its elevation and front/back hemifield). Several previous studies have shown that disrupting spectral cues by physically manipulating the pinnae reduces the accuracy of localisation in the up-to-down and front-to-back dimensions more so than in the left-to-right dimension [e.g., Roffler & Butler 1968a].

The effect of spectral scrambling on localisation accuracy was clearly greater for stimuli generated by randomising sound levels in broader spectral bands. The effect of spectral scrambling for the 98-Hz bandwidth was particularly small. This is probably a reflection of the spectral smoothing imposed on stimuli by the cochlea. The cochlea can be conceived of as a bank of bandpass filters tuned to different frequencies. The bandwidths of these filters increase with increasing centre frequency [e.g., Glassberg & Moore 1990]. Cochlear filters tuned to frequencies above 680 Hz (the centre frequency of the critical band, or filter, with a 98-Hz bandwidth) would have rejected much of the detail in stimuli generated by randomising sound levels in 98-Hz wide bands. This is because the bandwidths of these filters are greater than 98 Hz and the filters, as a consequence, would have integrated energy across multiple level-randomised bands. This would have had an effect equivalent to that of reducing the band-level ranges of the stimuli in the frequency range above 680 Hz. For stimuli generated by randomising sound levels in 391- and 1562-Hz wide bands, the frequencies above which detail would have been reduced by cochlear filtering are 3390 and 14240 Hz, respectively. For the latter bandwidth, therefore, almost all of the detail in the audible frequency range would have survived cochlear filtering.

As noted in the Results section, the effect of cochlear smoothing can be approximated by passing stimuli through a set of 1/3-octave bandpass filters. We have shown in this report that a simple measure of spectral variation based on the outputs of the eight 1/3-octave bandpass filters tuned to frequencies ranging from 2.5 to 12.5 kHz is highly predictive of the accuracies with which the elevations and front/back hemifields of the stimuli in this study could be discerned. It is likely that this measure will also be predictive of the accuracy with which similar stimuli (i.e., those with audible sound levels in most of the 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz and spectra that are more-or-less constant across time) can be localised. The extent to which this measure will be predictive of the accuracy with which stimuli with time-varying spectra can be localised, however, is less clear. Such stimuli include speech and other naturally occurring sounds that are being considered for use as spatialised threat warnings in military aviation environments. It is plausible that the accuracy with which many stimuli with time-varying spectra can be localised will be determined by the accuracy with which the most localisable portions of those stimuli can be localised. If that is the case, it should be possible to predict the accuracy with which stimuli with time-varying spectra can be localised by dividing them into segments of appropriate length and calculating the measure of spectral variation described here for each segment.

In summary, the study described here was conducted to enhance our understanding of the relationship between a sound's spectral variation and the accuracy with which it can be localised with a view to informing the development of metrics and guidelines for designing localisable auditory warnings. Our study has shown that the auditory localisation system is surprisingly tolerant of variation in a sound's spectrum provided that variation does not result in the sound having an extremely limited spectral range (i.e., provided the sound can be reasonably described as broadband). Nevertheless, severe spectral variation (i.e., spectral scrambling involving 40 and 60 dB band-level ranges) was found to result in a clear decrease in localisation accuracy, specifically with regard to judgments of sound-source elevation and front/back hemifield. It was found that the accuracy with which the spectrally scrambled stimuli in this study could be localised was

accurately predicted on the basis of a simple measure of spectral variation. It seems likely that this measure, the sum of the absolute differences between levels in adjacent 1/3-octave bands with centre frequencies ranging from 2.5 to 12.5 kHz, will also be predictive of the accuracy with which other sounds can be localised. Further research is required to determine whether that is the case.

5. ACKNOWLEDGMENTS

The authors wish to thank Bob Bolia of AFRL for providing helpful comments on an earlier version of this report. Some of the data presented here have been described previously in an honours thesis [Watt, 1997] and, as an abstract, in a conference proceedings [McAnally *et al.* 2000].

6. REFERENCES

Blauert, J. (1969/70). Sound localization in the median plane. *Acustica*, **22**, 205-213.

Bolia, R.S., D'Angelo, W.R., & McKinley, R.L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors*, **41**, 664-669.

Bronkhorst, A.W. (1995). Localization of real and virtual sound sources. *Journal of the Acoustical Society of America*, **98**, 2542-2553.

Bronkhorst, A.W., Veltman, J.A., & van Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. *Human Factors*, **38**, 23-33.

Burlinghame, J.A., & Butler, R.A. (1998). The effects of attenuation of frequency segments on binaural localization of sound. *Perception and Psychophysics,* **60**, 1374-1383.

Butler, R.A., & Musicant, A.D. (1993). Binaural localization: Influence of stimulus frequency and the linkage to covert peak areas. *Hearing Research,* **67**, 220-229.

Butler, R.A., & Planert, N. (1976). The influence of stimulus bandwidth on localization of sound in space. *Perception and Psychophysics*, **19**, 103-108.

Carlile, S., Leong, P., Pralong, D., Boden, R., & Hyams, S. (1996). High fidelity virtual auditory space: An operational definition. In S. Sestito, P. Beckett, G. Tudor, & T. Triggs, (Eds.) *Proceedings of Simtect 96*, 79-84, Melbourne, Simtect Organising Committee.

Corso, J.F. (1963). Age and sex differences in pure-tone thresholds. *Archives of Otolaryngology*, **77**, 385-405.

Glassberg, B.R., & Moore, B.C.J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, **47**, 103-138.

Hebrank, J., & Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane. *Journal of the Acoustical Society of America*, **56**, 1829-1834.

Hofman, P.M., & Van Opstal, A.J. (2002). Binaural weighting of pinna cues in human sound localization. *Experimental Brain Research*, **148**, 458-470.

Humanski, R.A., & Butler, R.A. (1988). The contribution of the near and far ear toward localization of sound in the sagittal plane. *Journal of the Acoustical Society of America*, **83**, 2300-2310.

Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook*, New Jersey, Prentice Hall.

King, R.B., & Oldfield, S.R. (1997). The impact of signal bandwidth on auditory localization: Implications for the design of three-dimensional audio displays. *Human Factors*, **39**, 287-295.

Macpherson, E.A., & Middlebrooks, J.C. (2003). Vertical-plane sound localization probed with ripple-spectrum noise. *Journal of the Acoustical Society of America,* **114**, 430-445.

Martin, R.L., McAnally, K.I., & Senova, M.A. (2001). Free-field equivalent localization of virtual audio. *Journal of the Audio Engineering Society*, **49**, 14-22.

McAnally, K.I., Martin, R.L., Watt, T., & Flanagan, P. (2000). The effect of spectral roughness on sound localisation (abstract). *Proceedings of the Australian Neuroscience Society,* **11***,* 145.

McKinley, R.L., Erikson, M.A., & D'Angelo, W.R. (1994). 3-Dimensional auditory displays: Development, applications, and performance. *Aviation, Space, and Environmental Medicine*, **65**, A31-38.

Mehrgardt, S., & Mellert, V. (1977). Transformation characteristics of the external human ear. *Journal of the Acoustical Society of America*, **61**, 1567-1576.

Middlebrooks, J.C. (1992). Narrow-band sound localization related to external ear acoustics. *Journal of the Acoustical society of America*, **92**, 2607-2624.

Middlebrooks, J.C., & Green, D.M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, **42**, 135-159.

Mills, A.W. (1972). Auditory localization. In J.V. Tobias, (Ed.) *Foundations of Modern Auditory Theory*, 303-348, New York, Academic.

Morimoto, M. (2001). The contribution of two ears to the perception of vertical angle in sagittal planes. *Journal of the Acoustical Society of America*, **109**, 1596-1603.

Musicant, A.D., & Butler, R.A. (1985). Influence of monaural spectral cues on binaural localization. *Journal of the Acoustical Society of America*, **77,** 202-208.

Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Brickman, B.J., Haas, M.W., & McKinley, R.L. (1998). Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Human Factors*, **40**, 452-460.

Parker, S.P.A., Smith, S.E., Stephan, K.L., Martin, R.L., & McAnally, K.I. (2004). Effects of supplementing head-down displays with 3-D audio during visual target acquisition. *International Journal of Aviation Psychology*, **14**, 277-295.

Perrott, D.R., Cisneros, J., McKinley, R.L., & D'Angelo, W.R. (1996). Aurally aided visual search under virtual and free-field listening conditions. *Human Factors*, **38**, 702-715.

Pratt, C.C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, **13**, 278-285.

Rakerd, B., Hartmann, W.M., & McCaskey, T. L. (1999). Identification and localization of sound sources in the median sagittal plane. *Journal of the Acoustical Society of America,* **106**, 2812-2820.

Roffler, S.K., & Butler, R.A. (1968a). Factors that influence the localization of sound in the vertical plane. *Journal of the Acoustical Society of America*, **43**, 1255-1259.

Roffler, S.K., & Butler, R.A. (1968b). Localization of tonal stimuli in the vertical plane. J*ournal of the Acoustical society of America*, **43**, 1260-1266.

Shaw, E.A.G., & Teranishi, R. (1968). Sound pressure generated in an external-ear replica and real human ears by a nearby point source. *Journal of the Acoustical Society of America*, **44**, 240-249.

Stelmachowicz, P.G., Beauchaine, K.A., Kalberer, A., & Jesteadt, W. (1989). Normative thresholds in the 8- to 20-kHz range as a function of age. *Journal of the Acoustical Society of America*, **86**, 1384-1391.

Watt, T. (1997). *The effect of frequency modifications on sound localization.* Unpublished honours thesis, Deakin University, Waurn Ponds, Australia.

Watson, D.B., Martin, R.L., McAnally, K.I., Smith, S.E., & Emonson, D.L. (2000). Effect of normobaric hypoxia on auditory sensitivity. *Aviation, Space, and Environmental Medicine*, **71**, 791-797.

Wightman, F.L., & Kistler, D.J. (1989). Headphone simulation of free-field listening. II: Psychophysical validation. *Journal of the Acoustical Society of America*, **85**, 868-878.

Wightman, F.L., & Kistler, D.J. (1997). Monaural sound localization revisited. *Journal of the Acoustical Society of America*, **101**, 1050-1063.

DISTRIBUTION LIST

The Effect of Spectral Variation on Sound Localisation

Russell Martin, Ken McAnally, Tavis Watt and Patrick Flanagan

AUSTRALIA

DEFENCE ORGANISATION

	No. of copies
Task Sponsor	
Chief, Air Operations Division	1 Printed
S&T Program	
Chief Defence Scientist	1
Deputy Chief Defence Scientist Policy	1
AS Science Corporate Management	1
Director General Science Policy Development	1
Counsellor Defence Science, London	Doc Data Sheet
Counsellor Defence Science, Washington	Doc Data Sheet
Scientific Adviser to MRDC, Thailand	Doc Data Sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	Doc Data Sht & Dist List
Scientific Adviser – Army	Doc Data Sht & Dist List
Air Force Scientific Adviser	Doc Data Sht & Dist List
Scientific Adviser to the DMO	Doc Data Sht & Dist List
Research Leader, CET, Air Operations Division	1 Printed
Head Training Research, Air Operations Division	1 Printed
Simon Parker, Air Operations Division	1 Printed
Author(s): Russell Martin, Air Operations Division	5 Printed
Ken McAnally, Air Operations Division	5 Printed
Tavis Watt, Deakin University	1 Printed
Patrick Flanagan, Deakin University	1 Printed
DSTO Library and Archives	
Library Fishermans Bend	1 Printed
Library Edinburgh	1 Printed
Defence Archives	1 Printed
Library, Sydney	Doc Data Sheet
Library, Stirling	Doc Data Sheet

Library Canberra	Doc Data Sheet
Capability Development Group	
Director General Maritime Development	Doc Data Sheet
Director General Capability and Plans	Doc Data Sheet
Assistant Secretary Investment Analysis	Doc Data Sheet
Director Capability Plans and Programming	Doc Data Sheet
Chief Information Officer Group	
Director General Australian Defence Simulation Office	Doc Data Sheet
AS Information Strategy and Futures	Doc Data Sheet
Director General Information Services	Doc Data Sheet
Strategy Group	
Director General Military Strategy	Doc Data Sheet
Assistant Secretary Governance and Counter-Proliferation	Doc Data Sheet
Navy	
Maritime Operational Analysis Centre, Building 89/90 Garden Island Sydney NSW Deputy Director (Operations) Deputy Director (Analysis)	Doc Data Sht & Dist List
Director General Navy Capability, Performance and Plans, Navy Headquarters	Doc Data Sheet
Director General Navy Strategic Policy and Futures, Navy Headquarters	Doc Data Sheet
Air Force	
SO (Science) - Headquarters Air Combat Group, RAAF Base, Williamtown NSW 2314	Doc Data Sht & Exec Summ
Army	
ABCA National Standardisation Officer Land Warfare Development Sector, Puckapunyal	e-mailed Doc Data Sheet
SO (Science) - Land Headquarters (LHQ), Victoria Barracks NSW	Doc Data & Exec Summary
SO (Science), Deployable Joint Force Headquarters (DJFHQ) (L), Enoggera QLD	Doc Data Sheet
Joint Operations Command	
Director General Joint Operations	Doc Data Sheet
Chief of Staff Headquarters Joint Operations Command	Doc Data Sheet
Commandant ADF Warfare Centre	Doc Data Sheet
Director General Strategic Logistics	Doc Data Sheet
COS Australian Defence College	Doc Data Sheet

Intelligence and Security Group	
AS Concepts, Capability and Resources	1
DGSTA , Defence Intelligence Organisation	1 Printed
Manager, Information Centre, Defence Intelligence Organisation	1
Director Advanced Capabilities	Doc Data Sheet
Defence Materiel Organisation	
Deputy CEO	Doc Data Sheet
Head Aerospace Systems Division	Doc Data Sheet
Head Maritime Systems Division	Doc Data Sheet
Program Manager Air Warfare Destroyer	Doc Data Sheet
CDR Joint Logistics Command	Doc Data Sheet
Guided Weapon & Explosive Ordnance Branch (GWEO)	Doc Data Sheet
OTHER ORGANISATIONS	
National Library of Australia	1
NASA (Canberra)	1
UNIVERSITIES AND COLLEGES	
Australian Defence Force Academy	
Library	1
Head of Aerospace and Mechanical Engineering	1
Hargrave Library, Monash University	Doc Data Sheet

OUTSIDE AUSTRALIA

INTERNATIONAL DEFENCE INFORMATION CENTRES	
US Defense Technical Information Center	1
UK Dstl Knowledge Services	1
Canada Defence Research Directorate R&D Knowledge & Information Management (DRDKIM)	1
NZ Defence Information Centre	1
ABSTRACTING AND INFORMATION ORGANISATIONS	
Library, Chemical Abstracts Reference Service	1
Engineering Societies Library, US	1
Materials Information, Cambridge Scientific Abstracts, US	1
Documents Librarian, The Center for Research Libraries, US	1
SPARES	5 Printed
Total number of copies: 44	**25 Printed** **19 PDF**

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
The Effect of Spectral Variation on Sound Localisation	Document (U) Title (U) Abstract (U)
4. AUTHOR(S)	5. CORPORATE AUTHOR
Russell Martin, Ken McAnally, Tavis Watt and Patrick Flanagan	DSTO Defence Science and Technology Organisation 506 Lorimer Street Fishermans Bend, Victoria 3207 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-RR-0308	AR-013-607	Research Report	March, 2006

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
2004/1030782/1	05/115	Chief AOD	16	38

13. URL on the World Wide Web	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-RR-0308.pdf	Chief, Air Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DSTO RESEARCH LIBRARY THESAURUS

Auditory localisation

19. ABSTRACT

Previous research has suggested that spatialised auditory displays will enhance operator performance in many military settings. It is well known that a sound's spectrum must be broad and relatively flat for the sound to be accurately localised. The study described here examined the effect of systematically varying the evenness of a sound's spectrum on the accuracy with which the sound can be localised. Six participants localised spectrally scrambled sounds produced by setting the sound levels in the 98-, 391- or 1562-Hz wide frequency bands comprising a broadband (0-25 kHz) sound to random values within a 0-, 20-, 40- or 60-dB range. Localisation errors were found to increase with increases in both bandwidth and band-level range. Scrambling the spectra of sounds over a 60 dB range led to as much as a doubling of mean elevation error and a trebling of front/back confusion rate. The accuracy with which these sounds could be localised was found to be highly correlated with a simple measure of spectral variation. The results of this study inform the development of guidelines for designing localisable sounds to be used in spatialised auditory displays.

Page classification: UNCLASSIFIED